基于改进实例分割算法的区域养殖生猪计数系统

张岩琪[1,2]，周 硕[1,2]，张 凝[1,2*]，柴秀娟[1,2]，孙 坦[1,2]

（1. 中国农业科学院 农业信息研究所，北京 100081，中国；2. 农业农村部农业大数据重点实验室，北京 100081，中国）

摘 要：[**目的/意义**] 针对现有规模化猪场生猪计数需求场景多，人工计数效率低、成本高等问题，提出一种基于改进实例分割深度学习算法和微信公众平台的区域养殖生猪计数方法。[**方法**] 首先，利用智能手机拍摄养殖场猪只视频，对视频抽帧进一步生成图像数据集。其次，通过改进卷积块注意力模块（Convolutional Block Attention Module, CBAM）中忽略通道与空间相互作用及通道注意力中降维操作带来的效率较低问题，提出高效全局注意力模块，并将该模块引入基于回归分析的单阶段实例分割网络 YOLO（You Only Look Once）v8 中对获取的生猪图像进行分割，构建新的识别模型 YOLOv8x-Ours，以实现高精度的生猪计数。最后，基于微信公众平台开发微信小程序，并嵌入综合表现最优的生猪计数模型，实现使用智能手机拍摄图像进行生猪快速计数。[**结果和讨论**] 在测试集上的试验结果表明，与现有实例分割模型相比，引入高效全局注意力的 YOLOv8x-Ous 模型获得 66% 的平均精度（$AP_{(50:95)}$），平均绝对误差（Mean Absolute Error, MAE）、均方根误差（Root Mean Square Error, RMSE）和 R^2 分别为 1.727、2.168 和 0.949，表现出较高的准确性和稳定性。模型计算猪只数量误差小于 3 头猪的图像数量占测试图像总数量的 93.8%，相比两阶段实例分割算法 Mask R-CNN（Region Convolutional Neural Network）提升 7.6%；单幅图像平均处理时间仅为 64 ms，是 Mask R-CNN 的 1/8。[**结论**] 该方法经济高效，为规模化猪场的生猪计数提供了一种技术方案。

关键词：生猪计数；深度学习；微信小程序；YOLOv8；实例分割

中图分类号：TP391.41　　**文献标志码：**A　　**文章编号：**SA202310001

引用格式：张岩琪, 周硕, 张凝, 柴秀娟, 孙坦. 基于改进实例分割算法的区域养殖生猪计数系统[J]. 智慧农业(中英文), 2024, 6(4): 53-63. DOI：10.12133/j.smartag.SA202310001

ZHANG Yanqi, ZHOU Shuo, ZHANG Ning, CHAI Xiujuan, SUN Tan. A Regional Farming Pig Counting System Based on Improved Instance Segmentation Algorithm[J]. Smart Agriculture, 2024, 6(4): 53-63. DOI：10.12133/j. smartag. SA202310001 (in Chinese with English abstract)

0 引 言

生猪养殖是畜牧业的重要组成部分。猪肉是国内人们最主要的动物蛋白来源，也是最主要的肉类食品，约占国民人均肉类消费量的60%。然而，受疾病和环保政策等影响，虽然目前生猪养殖模式以散养为主，但规模化、现代化养殖逐渐成为主趋势。根据舍内猪只数量进行精准化饲养和管理，对规模化养殖有着重要意义。

目前，养殖公司是以人工统计的方式进行出栏量和存栏量的资产盘点，费时费力，且往往由于猪只跑动导致错误计数并存在作弊问题。随着养殖规模的不断扩大，周期性的活体资产盘点工作导致养殖企业面临人力、物力、财力及时间紧迫的多重压力。采用电子耳标获取猪只信息可以实现生猪计数。但生猪在群养过程中耳标易坏易掉，导致计数结果不准确，并且耳标属于入侵式传感设备，易导致细菌感染，使得猪只产生较为强烈的应激反应。

随着深度学习技术的不断发展，卷积神经网络[1]、视觉自注意力模型（Vision Transformer,

收稿日期：2023-09-28

基金项目：新一代人工智能国家科技重大专项（2022ZD0115702）；国家自然科学基金项目（61976219）；北京市智慧农业创新团队项目资助（BAIC10-2024）；中国农业科学院创新工程（CAAS-ASTIP-2023-AII）；中央级公益性科研院所基本科研业务费专项（JBYW-AII-2023-04，JBYW-AII-2022-14）

作者简介：张岩琪，研究方向为计算机视觉与农业机器人。E-mail：zhangyanqi@caas.cn

*通信作者：张 凝，博士，副研究员，研究方向为作物表型与定量遥感。E-mail：zhangning@caas.cn

ViT）[2] 等深度学习算法在计算机视觉任务中表现优异。越来越多的研究将深度学习算法应用到农业和畜牧业盘点计数领域。通过人工智能计算猪只的方式，自动统计养殖场每天/每周的生猪出栏量和存栏量，在一定误差允许范围内可免去人工点数费时费力的问题，达到机器替代人工的目的，有利于养殖场加强对养殖过程的监管，减少人力投入和管理成本。Chen 等 [3] 提出了一种自底向上的猪只身体关键点检测方法，并基于空间感知时序响应对视频流中的猪只进行计数。Tian 等 [4] 根据 ResNeXt [5] 学习从图像特征到密度图的映射关系，通过整合密度图获得了图像中猪只的总数，对于每幅图像中平均有15头猪只的数据集获得了1.67的平均绝对误差。高云等 [6] 通过添加空间金字塔结构改进人群计数网络 GSRNet（Genomic Signals and Regions）[7]，有效提高了高密度猪只的计数精度。Kim 等 [8] 基于深度学习视频检测和跟踪方法对猪场走廊摄像头的视频流进行分析，实现了统计通过计数区的猪只数量，计数准确率达到99.44%。Miso 等 [9] 使用 YOLO（You Only Look Once）网络检测目标猪只，再结合 Kinect 深度相机实现猪只的实时分割。王荣等 [10] 融合多尺度特征金字塔网络和可变形卷积 [11]，解决猪只间粘连、遮挡等问题，在科大讯飞公开的生猪盘点数据集的分割准确率达到96.7%。Xu 等 [12] 使用 Mask R-CNN（Region Convolutional Neural Network）[13] 对四旋翼无人机拍摄的牛羊图像进行分割和盘点计数，对畜牧的分类准确率为96%，畜牧数量估计精度达到92%。Sarwar 等 [14] 使用全卷积神经网络分割无人机捕获的航空图像中的牲畜个体，相比 FCN（Full Connected Network），召回率从90%提升至98%。胡云鸽等 [15] 将 Mask R-CNN 网络的底层边缘位置特征与高层特征相融合，以提升网络对猪只边缘轮廓的识别能力。以上方法从俯视角度拍摄图像，需要在猪舍每个猪栏上方安装摄像设备或使用无人机进行拍摄，成本需求较大。

针对上述人工点数效率低、电子耳标伤害猪只且容易脱落、传统图像处理精度低和俯视角度采集图像成本大等现有生猪计数方法的问题，为贴合规模化猪场区域内生猪计数任务的需求，本研究利用深度学习实例分割算法和微信公众平台，提出一种高效率且低成本的生猪计数方法。主要工作包括：1）使用智能手机采集区域内生猪图像数据；2）在 YOLOv8 中引入本研究提出的高效全局注意力模块构建生猪计数实例分割网络，并通过实验发现生猪计数的最优深度学习模型；3）依托微信公众平台，开发“生猪计数”微信小程序，实现拍摄区域内生猪图像并计数；4）在小程序添加用户交互功能，对于图像中深度学习模型漏检错检的生猪，允许用户点击相应生猪位置进行增加和减少计数结果。

1 材料与方法

1.1 数据采集

于2022年10月在正邦集团有限公司旗下猪场进行数据采集。采集目标为猪舍中每个猪栏内的群养猪只。使用智能手机的后置摄像头模拟人的视角进行视频数据的采集，拍摄时摄像头位于拍摄人的面部前方。共采集视频数据70段，每段视频时间长度为8～92 s，分辨率为720×1 280像素，帧速率为30帧/s，使用MP4格式储存。

1.2 数据集构建

对采集的群养猪只视频数据，通过编写代码以15帧的间隔抽取视频中的图像，并剔除过度相似、模糊和两个栏位之间的图像以构建数据集。最终获得671幅图像，并按9∶1划分训练集（606幅）和测试集（65幅）。数据集在自然光及较暗的灯光场景下采集。图像中的猪只包括2～3月龄的保育猪和体型较大的育肥猪。猪只状态包含进食、站立、坐、趴卧、打闹及相互遮挡等。数据集的部分示例如图1所示。

使用Labelme标注工具，通过鼠标沿图像中每头猪只的轮廓框选出多边形进行标注。多边形的每个点以坐标点的形式储存，标注文件为JSON格式，每幅图像中多边形的个数为猪只的真实个数。共计标注13 655头猪只。单张图像中猪只计数范围4～63头。生猪计数数据集如表1所示。

1.3 生猪自动计数模型构建

本研究提出了一种使用智能手机拍摄区域养殖生猪图像的生猪计数方法。通过改进卷积块注意力模块（Convolutional Block Attention Module, CBAM）中忽略通道与空间相互作用及通道注意力中降维操作带来的效率较低问题，提出高效全局注意力模块，并将该注意力引入基于回归分析的单阶段实例分割网络 YOLOv8 中对获取的生猪图像进行检测，实现高精度的生猪计数。



图 1　生猪计数数据集示例

Fig. 1 Examples of pig counting dataset

表 1　改进实例分割算法的生猪计数数据集

Table 1 Pig count dataset created in improved instance segmentation algorithm

数据集	图像数量	猪只总数/头	单幅图像最小猪只数/头	单幅图像最大猪只数/头
训练集	606	11 609	4	63
测试集	65	2 046	10	46

1.3.1　YOLOv8

YOLOv8 是 Ultralytics 开发的目标检测和图像分割模型的最新版本[16]，建立在现有 YOLO 版本的基础上，引入新的主干网络、新的 Anchor-Free 检测头和新的损失函数，并能够在各种硬件平台上运行，其结构如图2所示。

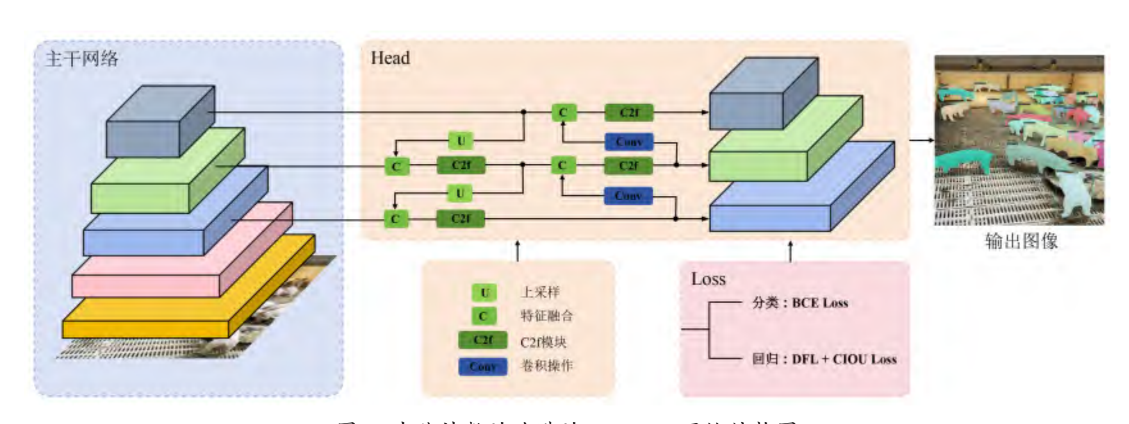


图 2　生猪计数的改进的 YOLOv8 网络结构图

Fig. 2 Structure of the improved YOLOv8 Network for pig counting

相比之前的版本，YOLOv8在主干网络和Head中使用C2f替换所有C3模块，增加更多的跳层连接和Split操作，实现进一步轻量化的同时获得更加丰富的梯度流信息。主干网络使用快速空间金字塔池化（Spatial Pyramid Polling-Fast, SPPF）模块进行特征提取，相比空间金字塔池化（Spatial Pyramid Polling, SPP）[17]，模型的计算量变小以提升计算速度。在 Head 模块中，YOLOv8 对检测和分类进行解耦，采用不同的分支进行运算以提升模型精度。同时使用 Anchor-Free，不需要预设 Anchor，只对不同尺度的特征图的目标中心点和宽高进行回归，极大地提升了运算速率。本研究中采用CIOU_Loss

(Complete Intersection over Union Loss)[18] 和非对称加权的 DFL_Loss (Distributional Feature Loss)作为回归的损失函数，增强对遮挡、重叠猪只的检测性能。

1.3.2 高效全局注意力模块

CBAM通过学习的方式获取每个特征通道和空间的重要程度，以此来提升特征并抑制对当前任务不重要的特征。CBAM的局限主要有：1）忽略了特征通道与空间的相互作用，从而丢失跨维信息。2）在通道注意力模块中，使用两层全连接的降维操作对通道注意力的预测产生负面影响，且计算成本高。对此，本研究提出了一种高效全局注意力模块，其结构如图3所示。

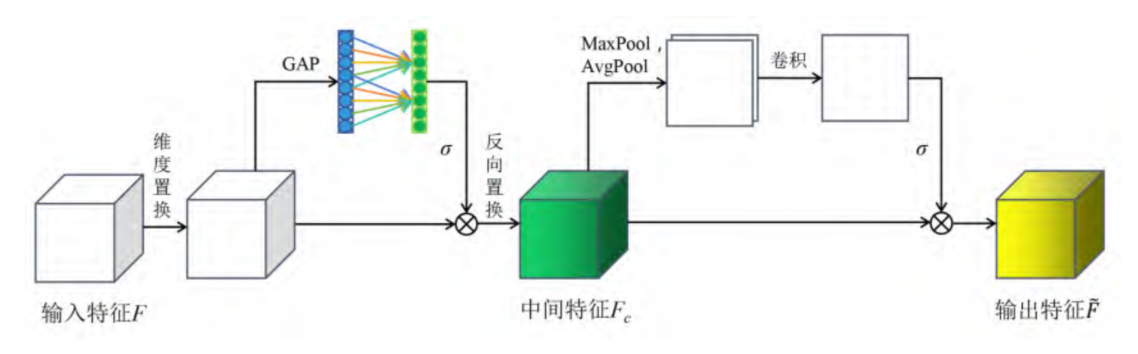


图3 高效全局注意力模块结构图

Fig. 3 Structure of efficient global attention module

高效全局注意力模块依次进行通道和空间注意力操作，在通道注意力子模块中，为获取输入特征$F \in R^{C \times W \times H}$通道与空间维度的相互作用，首先使用维度置换操作将输入特征的维度转换为$W \times H \times C$，对转换后的特征使用全局平均池化（Global Average Pooling, GAP）进行聚合，使用一维卷积代替CBAM中的全连接操作，避免降维的同时极大地降低了模型的参数量，再使用sigmoid激活函数并与特征进行对应位置乘积，最后执行反向置换得到中间特征F_c，通道注意力子模块可以表示为公式（1）。

$$F_c = \mathrm{permute}'\{\sigma(\mathrm{Conv1}D[\mathrm{GAP}(\mathrm{permute}(F))]) \otimes \mathrm{permute}(F)\} \quad (1)$$

式中：permute和permute'分别为维度置换和反向置换操作；σ为sigmoid激活函数。

空间注意力子模块与CBAM类似，对中间特征F_c在通道维度分别使用平均池化和最大池化，将产生的特征图进行拼接（concat），然后在拼接后的特征图上使用卷积操作产生最终的空间注意力图，与F_c进行对应位置乘积后得到输出特征$\tilde{F}$，空间注意力子模块表示为公式（2）。

$$\tilde{F} = \sigma\left\{\mathrm{Conv2}D\left(\begin{bmatrix}\mathrm{MaxPool}(F_c) \\ \mathrm{AvgPool}(F_c)\end{bmatrix}\right)\right\} \quad (2)$$

本研究通过将高效全局注意力模块添加至YOLOv8主干网络的每个C2f层来构建生猪计数网络模型，可以更加有效地提取图像中的维度依赖关系和特征信息，从而实现高精度的生猪计数。

1.4 生猪计数小程序开发

针对养殖场高效率、低成本的要求，本研究基于微信公众平台和Django Web框架开发生猪计数微信小程序，嵌入本研究提出的计数模型，通过微信平台使用智能手机拍摄图像实现生猪计数任务。微信小程序具有无需安装、用时随时打开、不占用手机空间、功能实现完全可以与App媲美等优点，并且背靠11亿微信用户，共享微信传播生态（公众号、服务号、订阅号、朋友圈等），从用户和开发角度出发都非常适合应对当前任务需求。微信小程序计数方法技术路线如图4所示，主要包含以下功能：

1）登录。用户进入生猪计数小程序，进行授权登录。

2）创建猪场结构。新用户根据猪场布局，创建猪仓-猪舍-猪栏3级结构。

3）获取图像。选择需要统计生猪数量的猪栏，对其进行拍照或从相册中选取对应的图像。

4）统计数量。将照片输入最优的生猪计数模型，根据模型的预测结果统计图像中生猪数量。

5）用户交互。生猪计数要求准确率高，对于模型存在错检漏检的情况，允许用户点击错误位置修改生猪数量。

6）历史记录。查看历史计数信息，包括栏位信息、检测时间、计数结果图像和计数结果。

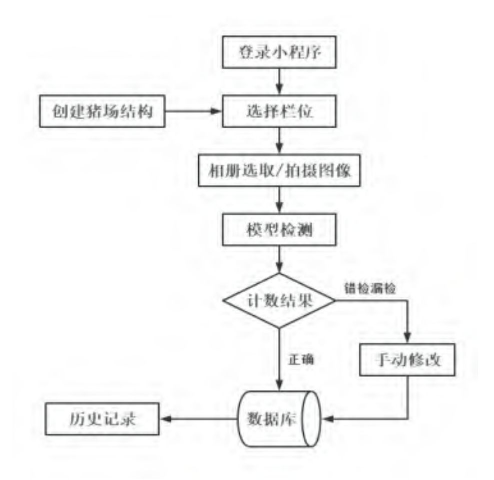


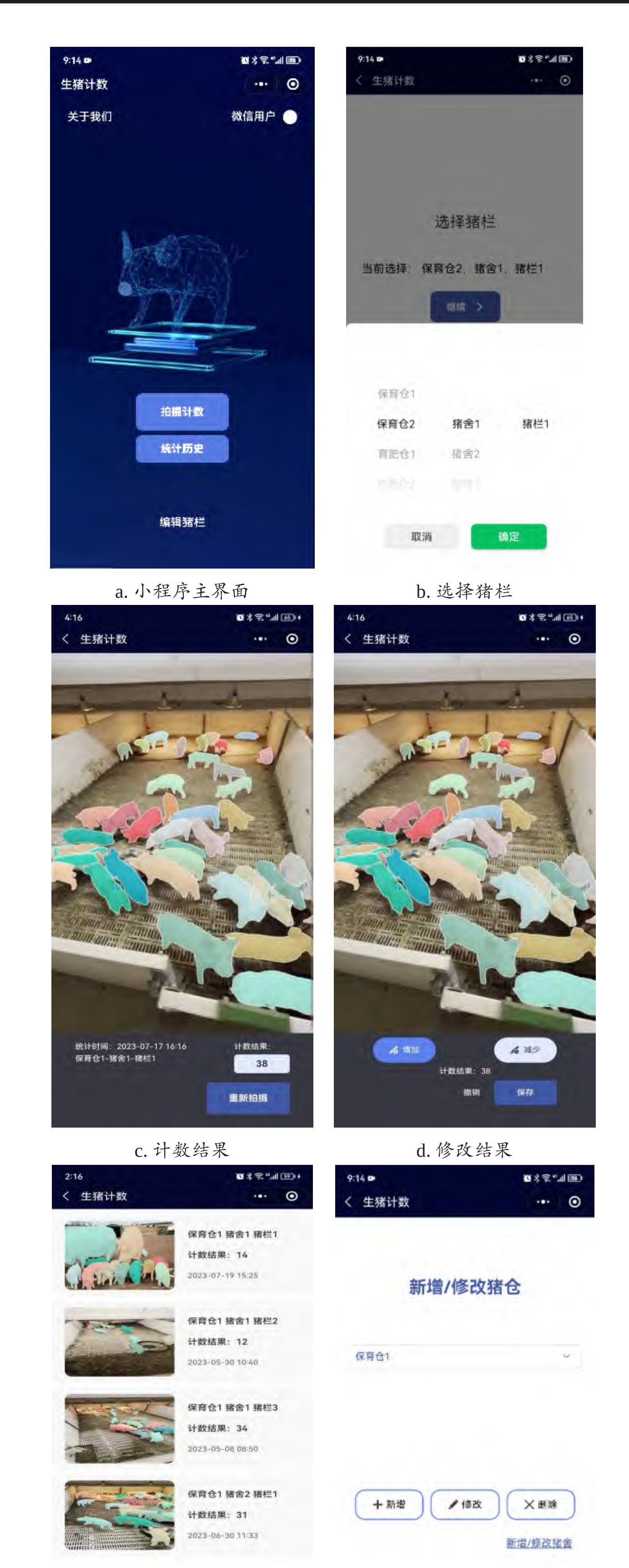
图4　生猪计数微信小程序设计技术路线图

Fig. 4 Technical route for pig counting WeChat mini program

生猪计数小程序的前端页面及页面间的跳转和数据传递功能使用微信公众平台的微信开发者工具进行开发。小程序对数据库的增删改查和拍摄图像的检测功能通过https接口的请求与响应实现，使用Django Web框架完成开发。图5为生猪计数小程序的主要操作界面。图5a为小程序的主界面，主要包括“用户信息的显示及拍摄计数、统计历史、编辑猪栏和关于我们”功能；点击拍摄计数按钮，用户需要选择进行计数的猪栏位置，如图5b所示；选择完成后进行区域内猪只图像拍摄，小程序调用智能手机的后置摄像头进行拍摄，获取的图像通过request方法的POST请求方式发送到后端代码，通过调用训练好的计数模型对图像中的猪只进行检测，最后得到图像分割和计数结果，如图5c所示，用户可以通过双指操作实现对分割图像的缩放，以便于更清楚地观察图像中小猪只和堆叠猪只的分割结果；对于存在错检漏检的统计，长按结果图像1 s进入修改结果界面，如图5d所示，点击模型分割出错位置修改相应的统计结果；点击统计历史按钮可以查看所有猪栏统计历史信息，如图5e所示；选择编辑猪栏按钮，对猪场的猪仓、猪舍和猪栏进行增删改查操作，如图5f所示。

1.5 实验环境配置

实验在Windows 10操作系统上进行，处理器为16核Intel（R）Core（TM） i9-9900K CPU @ 3.60 GHz，基于Pytorch 1.9.1框架搭建生猪计数深度学习算法，使用1块NVIDIA GeForce GTX



a. 小程序主界面　　b. 选择猪栏

c. 计数结果　　d. 修改结果

e. 历史记录　　f. 修改猪场结构

图5　生猪计数微信小程序主要操作界面

Fig. 5 Operation pages of pig counting WeChat mini program

2080Ti，11 GB显存GPU进行训练，CUDA版本为11.1，cuDNN版本为8.0.0。微信小程序前端使用微信开发者工具进行开发，后端基于Django3.2框架编写功能逻辑代码，使用MySQL数据库存储数据，版本为8.0.21。

网络训练使用SGD（Stochastic Gradient Descent）作为优化器，最大迭代次数100次，动量0.9，权重衰减为0.000 5，batchsize为8，输入图像尺寸为640×640。YOLOv8的初始学习率为0.01，Mask R-CNN的初始学习率为0.005。

1.6 评价指标

本研究从模型复杂度、推理时间和检测效果等方面量化分析生猪计数网络模型性能，主要包括模型参数量和平均每幅图像的推理时间。平均精度（Average Precision, AP）用于衡量对一个类检测的好坏，其结果是由精确率（Precision，P）和召回率（Recall，R）组成的PR曲线围成的面积。计算如公式（3）～公式（5）所示。

$$P=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \tag{3}$$

$$R=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \tag{4}$$

$$\mathrm{AP}=\int_0^1 P(R)dR \tag{5}$$

式中：TP（True Positive）为判断正样本正确的数量；FP（False Positive）为将负样本判断为正样本的数量；FN（False Negative）为将负样本判断错误的数量。根据IoU（Intersection over Union）取值不同，本研究报告了AP50和$\mathrm{AP}_{(50:95)}$，AP50为IoU的值取50%时AP结果。$\mathrm{AP}_{(50:95)}$是指IoU的值以5%的步长从50%取到95%，然后计算这些IoU下AP的均值。

为评估不同网络模型的生猪计数性能，选取平均绝对误差（Mean Absolute Error, MAE）、均方根误差（Root-Mean-Square Error, RMSE）和决定系数R^2作为计数效果的评价指标，对65幅生猪测试图像的预测结果进行评价。假设$\hat{Y}=\{\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_{65}\}$为预测值，$Y=\{y_1, y_2, \ldots, y_{65}\}$为真实值，$\bar{y}$为$Y$的均值，以上评价指标的计算如公式（6）～公式（8）所示。

$$\mathrm{MAE}=\frac{1}{65}\sum_{i=1}^{65}\left|\hat{y}_i-y_i\right| \tag{6}$$

$$\mathrm{RMSE}=\sqrt{\frac{1}{65}\sum_{i=1}^{65}\left(y_i-\hat{y}_i\right)^2} \tag{7}$$

$$R^2=1-\frac{\sum_{i=1}^{65}\left(y_i-\hat{y}_i\right)^2}{\sum_{i=1}^{65}\left(y_i-\bar{y}\right)^2} \tag{8}$$

式中：MAE为真实值与预测值绝对误差的平均值，表示模型预测的准确性，MAE值越小，模型的准确度越高；RMSE为真实值与预测值之差的样本标准差，表示模型预测的稳定性，RMSE值越小，模型的鲁棒性越高；决定系数R^2反映模型的拟合程度，取值范围为0～1，值越接近1，表示模型对数据的拟合能力越强。

2 结果与讨论

根据模型规模从小到大，YOLOv8提供YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l、YOLOv8x这5种网络结构。YOLOv8n是最小最快的模型；YOLOv8x是最准确但最慢的模型。根据生猪计数任务需求，本研究使用YOLOv8x作为基础模型，引入CBAM（YOLOv8x-CBAM）和高效全局注意力模块（YOLOv8x-Ours），并对比分析Mask R-CNN、YOLACT（Real-time Instance Segmentation）[19]、PolarMask[20]和SOLO（Segmenting Objects by Locations）[21]实例分割模型的性能。

2.1 不同计数模型性能对比分析

将不同实例分割模型的生猪计数网络在测试集上进行性能对比，结果如表2所示。

表2 不同实例分割模型性能比较

Table 2 Performance comparison of different instance segmentation models

网络模型	AP50	$\mathrm{AP}_{(50:95)}$	参数量/MB	单幅图像平均检测时间/ms
YOLACT	0.689	0.399	29.2	63
PolarMask	0.742	0.422	36.3	151
SOLO	0.811	0.537	46.2	113
Mask R-CNN	0.900	0.644	43.7	500
YOLOv5x	0.865	0.502	73.9	26
YOLOv8x	0.897	0.636	68.2	44
YOLOv8x-CBAM	0.901	0.646	125.3	109
YOLOv8x-Ours	0.901	0.660	71.7	64

在模型规模方面，YOLACT参数量最小，为29.2 MB，YOLOv8x-CBAM的参数量最大，约为YOLOv8x的两倍，而YOLOv8x-Ours相比YOLOv8x仅增加3.5 MB。YOLOv8x-Ours的AP50和

$AP_{(50:95)}$ 在所有模型中取得最好的性能，并且YOLOv8x-Ours处理图像的速度比Mask R-CNN快8倍。本研究使用的YOLOv5实例分割模型为参数规模最大的YOLOv5x，其处理图像的速度最快，平均每幅图像仅需26 ms，但平均精度在YOLO系列的模型中最低。综合来看，YOLOv8-Ours与其他模型相比具有最高的平均精度，且在处理图像速度方面有比较优异的表现。

2.2 不同计数模型结果对比分析

区域养殖的生猪图像通过网络模型预测区域内生猪的位置和掩膜。掩膜的个数即为图像中生猪的个数，通过设置分数阈值，删除目标得分低于该阈值的目标来控制图像中检测的生猪个数。表3为YOLOv8系列和其他实例分割模型在不同得分阈值时的预测结果。通过引入注意力机制，生猪计数模型的预测结果得到明显的提升。相比CBAM，本研究提出的高效全局注意力获得更好的表现，在得分阈值取0.15时，YOLOv8x-Ours获得最好的模型准确率和鲁棒性。在其他实例分割模型中，YOLACT预测的猪只数量少于真实值且得分均大于0.9，当得分阈值取较小数时，其结果保持不变。PolarMask、SOLO和Mask R-CNN分别在得分阈值为0.35、0.3和0.35时的预测结果最优。Mask R-CNN相比其他实例分割模型具有更好的表现。YOLOv8-Ours在所有模型中的准确率、稳定性和拟合能力表现最佳，比Mask R-CNN在得分阈值为0.35时的MAE、RMSE和 R^2 分别提升0.027、0.198和0.005。

表3 不同阈值下各个实例分割模型在测试集上的预测结果

Table 3 Prediction results of instance segmentation models on the test set with different thresholds

模型	评价指标	得分阈值							
		0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4
YOLACT	MAE	—				7.125	7.125	7.125	7.125
	RMSE					8.369	8.369	8.369	8.369
	R^2					0.327	0.327	0.327	0.327
PolarMask	MAE	—				5.512	5.578	2.853	4.157
	RMSE					7.132	6.596	3.775	5.018
	R^2					0.519	0.585	0.868	0.763
SOLO	MAE	—				3.637	2.913	3.141	4.085
	RMSE					4.792	3.634	4.066	4.978
	R^2					0.785	0.879	0.843	0.767
Mask R-CNN	MAE	—				1.969	1.769	1.754	1.769
	RMSE					2.649	2.421	2.366	2.376
	R^2					0.932	0.943	0.946	0.945
YOLOv8x	MAE	2.169	1.985	2.462	2.892	3.077	—		
	RMSE	2.826	2.715	3.153	3.526	3.713			
	R^2	0.922	0.928	0.903	0.879	0.866			
YOLOv8x -CBAM	MAE	1.985	1.831	1.923	1.835	2.431	—		
	RMSE	2.720	2.434	2.508	2.440	3.103			
	R^2	0.928	0.942	0.939	0.939	0.906			
YOLOv8x -Ours	MAE	1.969	1.798	1.727	1.799	2.316	—		
	RMSE	2.651	2.375	2.168	2.380	3.005			
	R^2	0.931	0.945	0.949	0.944	0.938			

注：“—”表示各个模型在给出的得分阈值内已找到最大值，其他阈值无需再计算对应的MAE、RMSE和 R^2。

生猪计数结果如表4所示。本研究分别统计了各个模型在65幅测试集图像上准确计数、误差小于2头猪、误差小于3头猪和误差大于2头猪的图像的数量和占比。本研究模型准确计数的图像数量为43幅，远高于其他模型，相较于YOLOv8x，准确计数的图像数量增加12.4%。Mask R-CNN准确

计数的图像数量优于YOLOv5x和YOLOv8x，但三者误差小于2头猪和3头猪的图像数量相当。引入CBAM注意力机制的YOLOv8x-CBAM模型全面提升了YOLOv8x的计数精度，但本模型在计数误差为0、小于2头猪和小于3头猪的图像占比表现最优，误差大于2头猪的图像仅有4幅，占测试集图像的6.2%。相较于表4中的其他模型，本模型实现了区域养殖生猪的高精度盘点计数。

表4 不同实例分割模型在测试集上的生猪计数结果

Table 4 Pig cuouting results for different instance segmentation models on the test set

模型	准确计数的图像		误差小于2头猪的图像		误差小于3头猪的图像		误差大于2头猪的图像	
	数量	占比/%	数量	占比/%	数量	占比/%	数量	占比/%
YOLCAT	22	33.8	40	61.5	52	80.0	13	20.0
YOLOv5x	29	44.6	48	73.8	56	86.2	9	13.8
PolarMask	33	50.8	47	72.3	56	86.2	9	13.8
SOLO	35	53.8	50	76.9	58	89.2	7	10.8
YOLOv8x	35	53.8	49	75.4	57	87.7	8	12.3
Mask R-CNN	37	56.9	49	75.4	56	86.2	9	13.8
YOLOv8x-CBAM	41	63.1	51	78.5	59	90.8	6	9.2
YOLOv8x-Ours	43	66.2	57	87.7	61	93.8	4	6.2

2.3 结果可视化

在微信小程序中，通过模型检测得到生猪个体的位置及掩膜，为方便观察模型预测结果，对检测到的每只生猪掩膜使用不同颜色显示在原始输入图像。生猪图像检测结果可视化如图6所示。图6a中模型正确统计了原图中的每只生猪；图6b中模型出现错检情况，在对图像数据进行人工标注时，由于猪只之间的遮挡、图像边缘不完整的现象，不可避免地会将猪只的一部分标注为待分割的目标类别，即1条猪腿、1只猪耳甚至1条猪尾巴都会被给予与完整猪只相同的监督信号，导致计数结果多1只，在微信小程序修改结果界面选择减少按钮，点击错误位置以修正计数结果；图6c中模型漏检了原图中左上角的一只猪，受限于本数据集中图像数量和质量方面的欠缺，有些数据样本即使人眼也难以分辨。所以模型在面对一些失焦、光线条件差、背景复杂的难例（如图6c中栏位远端的灯下区域）时会产生错检和漏检等问题，在微信小程序修改结果界面选择增加按钮，点击漏检位置以修正计数结果。

2.4 模型泛化能力比较

为进一步验证本模型的生猪计数能力，本研究从互联网下载17幅群养猪只图像及3幅只包含牛的图像进行生猪计数测试。计数结果如表5所示。完全计数正确的图像共13幅，其中在3幅牛的图像中，计数结果为0，即模型并不会将牛识别为猪只进行计数，仅有1幅图像中猪只计数的误差为3只猪。图7展示了网络图像在小程序中的计数结果，验证了本模型具有较好的泛化能力以及稳定性。

3 结　论

本研究针对规模化猪场区域养殖生猪计数问题，通过在单阶段实例分割算法YOLOv8中引入本研究提出的高效全局注意力模块构建生猪计数模型，实现对区域内的生猪计数。结果表明：本研究提出的生猪计数模型的准确率和稳定性方面表现最优，在测试集上，本模型在准确计数、误差小于2头猪和误差小于3头猪的图像占比均最高，计数误差小于3头猪的图像数量占总测试图像的93.8%，相比两阶段实例分割算法Mask R-CNN和引入CBAM注意力机制的YOLOv8x模型分别提升7.6%和3%。并且由于YOLOv8模型的单阶段检测和Anchor-Free结构，单幅图像的处理速度为64 ms，是Mask R-CNN的1/8。综合实验分析，为满足高效率、低成本的任务需求，通过将在测试集上表现最优的模型嵌入开发的微信小程序，实现使用智能手机拍摄区域内猪只图像实现生猪计数。对于模型错检漏检的猪只，允许用户点击结果图像中错误位置修改统计结果，以进一步提升生猪计数的准确性。

(a)原图1 (b)结果图1

a. 计数正确

(a)原图2 (b)结果图2 (c)修改图2

b. 计数误差为1(错检,红色框处)

(a)原图3 (b)结果图3 (c)修改图3

c. 计数误差为1(漏检,红色框处)

图6 微信小程序实现生猪计数结果可视化

Fig. 6 Visualization of pig counting results in WeChat mini program

表5 YOLOv8x-Ours在网络图像上的计数结果

Table 5 Counting results of YOLOv8x-Ours on network images

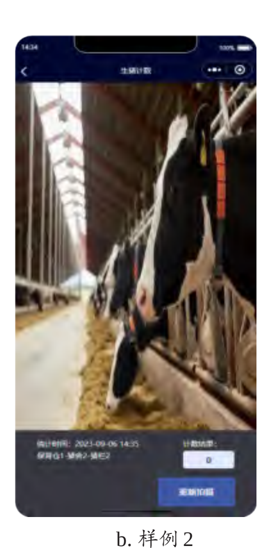
模型	准确计数的图像数量	误差小于2头猪的图像数量	误差小于3头猪的图像数量	误差大于2头猪的图像数量
YOLOv8-Ours	13	17	19	1

相比现有生猪计数方法，本研究无需在猪场的养殖区域上方架设众多成像设备，只需使用智能手机即可进行生猪计数，且用户可以通过图像分割结果快速辨别技术结果是否存在错误，对于错误结果可便捷地进行修改，使得本研究的方法更易于推广应用。这种人机协作的模式既节省了大量的人力，也可以保证计数结果的可靠性。

利益冲突声明： 本研究不存在研究者以及与公开研究成果有关的利益冲突。



a. 样例1



b. 样例2

图7 网络图像在生猪计数微信小程序中的计数结果

Fig. 7 Counting results of web images in the pig counting WeChat mini program

参考文献:

[1] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]// Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1. New York, USA: ACM, 2012: 1097-1105.

[2] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[EB/OL]. arXiv: 2010.11929, 2020.

[3] CHEN G, SHEN S W, WEN L Y, et al. Efficient Pig Counting in Crowds with Keypoints Tracking and Spatial-aware Temporal Response Filtering[C]// 2020 IEEE International Conference on Robotics and Automation (ICRA). Piscataway, New Jersey, USA: IEEE, 2020: 10052-10058.

[4] TIAN M X, GUO H, CHEN H, et al. Automated pig counting using deep learning[J]. Computers and electronics in agriculture, 2019, 163: ID 104840.

[5] XIE S N, GIRSHICK R, DOLLAR P, et al. Aggregated residual transformations for deep neural networks[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, New Jersey, USA: IEEE, 2017: 5987-5995.

[6] 高云, 李静, 余梅, 等. 基于多尺度感知的高密度猪只计数网络研究[J]. 农业机械学报, 2021, 52(9): 172-178.
GAO Y, LI J, YU M, et al. High-density pig counting net based on multi-scale aware[J]. Transactions of the Chinese society for agricultural machinery, 2021, 52(9): 172-178.

[7] LI Y H, ZHANG X F, CHEN D M. CSRNet: dilated convolutional neural networks for understanding the highly congested scenes[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, New Jersey, USA: IEEE, 2018: 1091-1100.

[8] KIM J, SUH Y, LEE J, et al. EmbeddedPigCount: Pig counting with video object detection and tracking on an embedded board[J]. Sensors, 2022, 22(7): ID 2689.

[9] JU M, CHOI Y, SEO J, et al. A kinect-based segmentation of touching-pigs for real-time monitoring[J]. Sensors, 2018, 18(6): ID 1746.

[10] 王荣, 高荣华, 李奇峰, 等. 融合特征金字塔与可变形卷积的高密度群养猪计数方法[J]. 农业机械学报, 2022, 53(10): 252-260.
WANG R, GAO R H, LI Q F, et al. High-density pig herd counting method combined with feature pyramid and deformable convolution[J]. Transactions of the Chinese society for agricultural machinery, 2022, 53(10): 252-260.

[11] ZHU X Z, HU H, LIN S, et al. Deformable ConvNets V2: more deformable, better results[C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, New Jersey, USA: IEEE, 2019: 9300-9308.

[12] XU B B, WANG W S, FALZON G, et al. Livestock classification and counting in quadcopter aerial images using Mask R-CNN[J]. International journal of remote sensing, 2020, 41(21): 8121-8142.

[13] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[J]. IEEE transactions on pattern analysis and machine intelligence, 2020, 42(2): 386-397.

[14] SARWAR F, GRIFFIN A, REHMAN S U, et al. Detecting sheep in UAV images[J]. Computers and electronics in agriculture, 2021, 187: ID 106219.

[15] 胡云鸽, 苍岩, 乔玉龙. 基于改进实例分割算法的智能猪只盘点系统设计[J]. 农业工程学报, 2020, 36(19): 177-183.
HU Y G, CANG Y, QIAO Y L. Design of intelligent pig counting system based on improved instance segmentation algorithm[J]. Transactions of the Chinese society of agricultural engineering, 2020, 36(19): 177-183.

[16] ULTRALYTICS. YOLOv8[EB/OL]. [2023-12-02]. https://github.com/ultralytics/ultralytics.

[17] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 37(9): 1904-1916.

[18] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[J]. Proceedings of the AAAI conference on artificial intelligence, 2020, 34(7): 12993-13000.

[19] BOLYA D, ZHOU C, XIAO F Y, et al. YOLACT: real-time instance segmentation[C]// 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, New Jersey, USA: IEEE, 2019: 9157-9166.

[20] XIE E Z, SUN P Z, SONG X G, et al. PolarMask: single shot instance segmentation with polar representation[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, New Jersey, USA: IEEE, 2020: 12193-12202.

[21] WANG X L, KONG T, SHEN C H, et al. SOLO: segmenting objects by locations[M]// VEDALDI A, BISCHOF H, BROX T, et al., Eds. Computer vision-ECCV 2020. Cham: Springer International Publishing, 2020: 649-665.

A Regional Farming Pig Counting System Based on Improved Instance Segmentation Algorithm

ZHANG Yanqi[1,2], ZHOU Shuo[1,2], ZHANG Ning[1,2*], CHAI Xiujuan[1,2], SUN Tan[1,2]

(1. *Agricultural Information Institute, Chinese Academy of Agricultural Sciences, Beijing* 100081, *China*;
2. *Key Laboratory of Agriculture and Rural Affairs, Beijing* 100081, *China*)

Abstract:

[Objective] Currently, pig farming facilities mainly rely on manual counting for tracking slaughtered and stored pigs. This is not only time-consuming and labor-intensive, but also prone to counting errors due to pig movement and potential cheating. As breeding opera-

tions expand, the periodic live asset inventories put significant strain on human, material and financial resources. Although methods based on electronic ear tags can assist in pig counting, these ear tags are easy to break and fall off in group housing environments. Most of the existing methods for counting pigs based on computer vision require capturing images from a top-down perspective, necessitating the installation of cameras above each hogpen or even the use of drones, resulting in high installation and maintenance costs. To address the above challenges faced in the group pig counting task, a high-efficiency and low-cost pig counting method was proposed based on improved instance segmentation algorithm and WeChat public platform.

[Methods] Firstly, a smartphone was used to collect pig image data in the area from a human view perspective, and each pig's outline in the image was annotated to establish a pig count dataset. The training set contains 606 images and the test set contains 65 images. Secondly, an efficient global attention module was proposed by improving convolutional block attention module (CBAM). The efficient global attention module first performed a dimension permutation operation on the input feature map to obtain the interaction between its channels and spatial dimensions. The permuted features were aggregated using global average pooling (GAP). One-dimensional convolution replaced the fully connected operation in CBAM, eliminating dimensionality reduction and significantly reducing the model's parameter number. This module was integrated into the YOLOv8 single-stage instance segmentation network to build the pig counting model YOLOv8x-Ours. By adding an efficient global attention module into each C2f layer of the YOLOv8 backbone network, the dimensional dependencies and feature information in the image could be extracted more effectively, thereby achieving high-accuracy pig counting. Lastly, with a focus on user experience and outreach, a pig counting WeChat mini program was developed based on the WeChat public platform and Django Web framework. The counting model was deployed to count pigs using images captured by smartphones.

[Results and Discussions] Compared with existing methods of Mask R-CNN, YOLACT(Real-time Instance Segmentation), PolarMask, SOLO and YOLOv5x, the proposed pig counting model YOLOv8x-Ours exhibited superior performance in terms of accuracy and stability. Notably, YOLOv8x-Ours achieved the highest accuracy in counting, with errors of less than 2 and 3 pigs on the test set. Specifically, 93.8% of the total test images had counting errors of less than 3 pigs. Compared with the two-stage instance segmentation algorithm Mask R-CNN and the YOLOv8x model that applies the CBAM attention mechanism, YOLOv8x-Ours showed performance improvements of 7.6% and 3%, respectively. And due to the single-stage design and anchor-free architecture of the YOLOv8 model, the processing speed of a single image was only 64 ms, 1/8 of Mask R-CNN. By embedding the model into the WeChat mini program platform, pig counting was conducted using smartphone images. In cases where the model incorrectly detected pigs, users were given the option to click on the erroneous location in the result image to adjust the statistical outcomes, thereby enhancing the accuracy of pig counting.

[Conclusions] The feasibility of deep learning technology in the task of pig counting was demonstrated. The proposed method eliminates the need for installing hardware equipment in the breeding area of the pig farm, enabling pig counting to be carried out effortlessly using just a smartphone. Users can promptly spot any errors in the counting results through image segmentation visualization and easily rectify any inaccuracies. This collaborative human-machine model not only reduces the need for extensive manpower but also guarantees the precision and user-friendliness of the counting outcomes.

Key words: pig counting; deep learning; WeChat mini program; YOLOv8; instance segmentation

Foundation items: National Science and Technology Major Project (2022ZD0115702); National Natural Science Foundation of China (61976219); Beijing Smart Agriculture Innovation Consortium Project (BAIC10-2024); Innovation Program of Chinese Academy of Agricultural Sciences (CAAS-ASTIP-2023-AII); Central Public-interest Scientific Institution Basal Research Fund (JBYW-AII-2023-04, JBYW-AII-2022-14)

Biography: ZHANG Yanqi, E-mail: zhangyanqi@caas.cn

***Corresponding author:** ZHANG Ning, E-mail: zhangning@caas.cn